

图文 90 设计一套Kafka到RocketMQ的双写+双读技术方案，实现无缝迁移！

413 人次阅读 2020-02-05 07:53:49

详情 评论



狸猫技术

进店逛

相关频道



从 0 开
间件实站
已更新1



继《从零开始带你成为JVM实战高手》后，救火队长携新作再度出山，重磅推荐：

(点击下方蓝字试听)

[《从零开始带你成为MySQL实战优化高手》](#)

今天的第二个技术问题，假设你们公司本来线上的MQ用的主要是Kafka，现在要从Kafka迁移到RocketMQ去，那么这个迁移的过程应该怎么做呢？应该采用什么样的技术方案来做迁移呢？

这里我们给大家介绍一个MQ集群迁移过程中的双写+双读技术方案。

简单来说，如果你要做MQ集群迁移，是不可能那么的简单粗暴的，因为你不可能说在某一个时间点突然之间就说把所有的Producer系统都停机，然后更新他的代码，接着全部重新上线，然后所有Producer系统都把消息写入到RocketMQ去了

一般来说，首先你要做到双写，也就是说，在你所有的Producer系统中，要引入一个双写的代码，让他同时往Kafka和RocketMQ中去写入消息，然后多写几天，起码双写要持续个1周左右，因为MQ一般都是实时数据，里面数据也就最多保留一周。

当你的双写持续一周过后，你会发现你的Kafka和RocketMQ里的数据看起来是几乎一模一样了，因为MQ反正也就保留最近几天的数据，当你双写持续超过一周过后，你会发现Kafka和RocketMQ里的数据几乎一模一样了。

但是光是双写还是不够的，还需要同时进行双读，也就是说在你双写的时候，你所有的Consumer系统都需要同时从Kafka和RocketMQ里获取消息，分别都用一模一样的逻辑处理一遍。

只不过从Kafka里获取到的消息还是走核心逻辑去处理，然后可以落入数据库或者是别的存储什么的，但是对于RocketMQ里获取到的消息，你可以用一样的逻辑处理，但是不能把处理结果具体的落入数据库之类的地方。

你的Consumer系统在同时从Kafka和RocketMQ进行消息读取的时候，你需要统计每个MQ当日读取和处理的消息的数量，这点非常重要，同时对于RocketMQ读取到的消息处理之后的结果，可以写入一个临时的存储中。

同时你要观察一段时间，当你发现持续双写和双读一段时间之后，如果所有的Consumer系统通过对比发现，从Kafka和RocketMQ读取和处理的消息数量一致，同时处理之后得到的结果也都是是一致的，此时就可以判断说当前Kafka和RocketMQ里的消息是一致的，而且计算出来的结果也都是是一致的。

这个时候就可以实施正式的切换了，你可以停机Producer系统，再重新修改后上线，全部修改为仅仅写RocketMQ，这个时候他数据不会丢，因为之前已经双写了一段时间了，然后所有的Consumer系统可以全部下线后修改代码再上线，全部基于RocketMQ来获取消息，计算和处理，结果写入存储中。

基本上对于类似的一些重要中间件的迁移，往往都会采取双写的方法，双写一段时间，然后观察两个方案的结果都一致了，你再正式下线旧的一套东西。

End

专栏版权归公众号**狸猫技术窝**所有

未经许可不得传播，如有侵权将追究法律责任

狸猫技术窝精品专栏及课程推荐：

[《从零开始带你成为JVM实战高手》](#)
[《21天互联网Java进阶面试训练营》（分布式篇）](#)
[《互联网Java工程师面试突击》（第1季）](#)
[《互联网Java工程师面试突击》（第3季）](#)

重要说明：

如何提问：每篇文章都有评论区，大家可以尽情留言提问，我会逐一答疑

如何加群：购买狸猫技术窝专栏的小伙伴都可以加入狸猫技术交流群，一个非常纯粹的技术交流的地方

具体加群方式，请参见目录菜单下的文档：《付费用户如何加群》（**购买后可见**）

